AFHRL-TR-76-10

12

FG

AIR FORCE

HUMAN RESOURCES

AD A023634

VISUAL SCANNING:
COMPARISONS BETWEEN STUDENT AND INSTRUCTOR PILOTS

By

Joseph DeMaio
Stanley Parkinson
Barry Leshowitz
John Crosby
Arizona State University
Tempe, Arizona 85281

Jack A. Thorpe, Capt, USAF

FLYING TRAINING DIVISION
Williams Air Force Base, Arizona 85224

June 1976
Final Report for Period 1 December 1974 – 1 December 1975

LABORATORY

AIR FORCE SYSTEMS COMMAND
BROOKS AIR FORCE BASE, TEXAS 78235

NOTICE

When US Government drawings, specifications, or other data are used for any purpose other than a definitely related Government procurement operation, the Government thereby incurs no responsibility nor any obligation whatsoever, and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data is not to be regarded by implication or otherwise, as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

This final report was submitted by Arizona State University, Tempe, Arizona 85281, under contract F41609-75-C-0018, project 2313, with Flying Training Division, Air Force Human Resources Laboratory (AFSC), Williams Air Force Base, Arizona 85224. Captain Jack A. Thorpe, Training Innovations Branch, was the contract monitor.

This report has been reviewed and cleared for open publication and/or public release by the appropriate Office of Information (OI) in accordance with AFR 190-17 and DoDD 5230.9. There is no objection to unlimited distribution of this report to the public at large, or by DDC to the National Technical Information Service (NTIS).

This technical report has been reviewed and is approved.

J. D. BOREN, Colonel, USAF
Chief, Flying Training Division

HAROLD E. FISCHER, Colonel, USAF
Commander

A

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

19

1. REPORT NUMBER: 18 AFHRL-TR-76-10	2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): 6 VISUAL SCANNING: COMPARISONS BETWEEN STUDENT AND INSTRUCTOR PILOTS,	5. TYPE OF REPORT & PERIOD COVERED: 9 Final rept. 1 Dec[redacted]974 – 1 Dec[redacted]975
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): 10 Joseph DeMaio, Stanley Parkinson, Barry Leshowitz, John Crosby, Jack A. Thorpe	8. CONTRACT OR GRANT NUMBER(s): 15 F41609-75-C-0018 NEW
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Arizona State University, Tempe, Arizona 85281	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 61102F; 16 AF-23130902; 17 231309
11. CONTROLLING OFFICE NAME AND ADDRESS: Hq Air Force Human Resources Laboratory (AFSC), Brooks Air Force Base, Texas 78235	12. REPORT DATE: 11 Jun[redacted]976
	13. NUMBER OF PAGES: 12 34p.
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): Flying Training Division, Air Force Human Resources Laboratory, Williams Air Force Base, Arizona 85224	15. SECURITY CLASS. (of this report): Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

visual scanning
instrument flying
cross-check
information processing
cognitive processing

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The performance of instructor pilots and student pilots was compared in two visual scanning tasks. In the first task both groups were shown slides of T-37 instrument displays. Some slides contained a significant deviation from a pre-determined straight and level course, and the task was to detect the error as quickly as possible. Instructor pilots detected errors faster and with greater accuracy than student pilots, thus providing evidence for the validity of the procedures employed. However, contrary to the concept of a fixed cross-check, student pilots showed a greater tendency to employ a systematic search pattern than did instructor pilots. This result suggests that rather than using a rigid scanning pattern, instructor pilots, by virtue of their additional flight experience, use a flexible scanning strategy which allows them to emphasize important or difficult aspects of the display. In the second experiment the

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

033 750 Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

next pg

mt

Item 20 (Continued)

attention diagnostic method task was employed to determine if the experience in visual scanning obtained in the flight situation would transfer to a novel scanning task. In the first session there were no differences in response latency between instructor pilots, student pilots, and a group of university students. Instructor pilots, however, showed a significant linear decrease in latency over the course of eight sessions while this trend was absent in the other two groups. This suggests that instructor pilots learn to attend to critical features more efficiently than do individuals with little or no flight experience. The results of the present experiments recommend the use of a variety of scanning tasks in the UPT program to facilitate the more rapid development of adaptive scanning strategies.

PREFACE

This report represents a portion of the research program of Project 2313, Perceptual-Motor and Cognitive Components of the Flying Task; Task 231309, Cognitive Components of the Flying Task, Dr Edward E. Eddowes, Task Scientist.

The authors are indebted to Captain Gale Webb, 82 FTW/DOR, Williams Air Force Base, Arizona, who juggled the busy flight line schedule of a number of pilots to coincide with the research needs of the experimenters. Sgt Michael Fiero of the Flying Training Division was also invaluable in the arranging of a smooth flow of experimental sessions. Dr Bernell Edwards offered an ideal room in the audiovisual laboratory for conduct of the experiment even though it caused some confusion in his production schedule.

PRECEDING PAGE BLANK-NOT FILMED

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

I. INTRODUCTION

The advanced aircraft of today is a highly sophisticated instrument, with numerous pilot aids, such as computerized automatic flight instruments and flight control systems. In many situations the physical control of the aircraft is of less importance than the pilot's ability to process information rapidly and accurately and to make correct decisions in real time. The pilot must remember a great deal of information entering through the visual, auditory and kinesthetic channels. This information is constantly being updated, and its criticality is subject to radical shifts at various points in the flight regime. For example, pilots must make visual judgments of distance, speed, and position under conditions which are dynamically changing and therefore predisposing to error.

One of the most dramatic changes, affecting pilot training since World War II, has been the shift from pure contact flying to the systematic use of aircraft instruments to achieve precision aircraft control. Today the use of instruments is the principal method of aircraft control regardless of weather conditions.

In 1932, the first system for aircraft control by instruments was introduced to train airline pilots under simulated flight conditions. This so-called "1-2-3 system of aircraft control" (AFM 51-37) consisted of three phases: (a) center the turn needle with the rudder, (b) center the ball with the aileron control, and (c) control the airspeed with the elevator. Despite the fact that this system offered a way of achieving straight and level flight during restricted visibility, it was not coordinated and hence difficult to employ with satisfactory results, and it did not take advantage of the pilot's natural ability to fly by visual reference.

It was not until the mid 1940's following the development of precision flight instruments that the current method for aircraft control, Instrument Flying, (AFM 51-37) was introduced.

The "control and performance concept" of attitude instrument flying refers to a set of procedural steps for manipulation and monitoring control and performance instruments. These procedural steps, as outlined in AFM 51-37, are as follows:

(1) Establish an attitude and/or power setting on the control instruments which should result in the desired

performance.

(2) Trim until control pressures are neutralized.

(3) Cross-check the performance instruments to determine if the established attitude and/or power setting are providing the desired performance.

(4) Adjust the attitude and power setting on the control instruments if a correction is necessary.

The present research is focused on the cross-check procedure in the T-37 instrument display. During the instrument flying phase of T-37 training (20.8 hrs.), pilot trainees receive cross-check demonstrations for each instrument maneuver. Cross-check instructions emphasize that attention must be divided between the control and performance instruments in a sequence that insures comprehensive coverage of the flight instruments. One cross-check or scanning strategy which is used by instructor pilots in training can be compared to a wagon wheel with the attitude indicator (AI) as the hub and the performance instruments as the spokes. Students are told that the AI is the only instrument that they must observe continuously for any length of time. During a given maneuver the student is instructed to start his cross-check or instrument scan with the AI. He monitors the AI, shifts to a performance instrument, returns to the AI, and then goes to another performance instrument. The cross-check proceeds in this fashion, including other performance instruments but always returns to the hub between adjacent performance instruments in the pattern. For example, consider straight and level unaccelerated flight which consists of maintaining a desired altitude, heading, and airspeed. A wagon-wheel cross-check might be AI, airspeed, AI, altimeter, AI, heading, etc.

Despite the importance of the cross-check technique in instrument flying, instructions for cross-checking emphasize flexibility and idiosyncracy in lieu of an optimal scanning pattern. While trainees do receive cross-check demonstrations for each maneuver as noted above, instructor pilots (IPs) are told to emphasize that such demonstrations are only examples of possible scanning strategies, and students are encouraged to develop cross-checks that "work best for them."

Cross-check strategies are thus developed by each pilot as a result of his experience in the aircraft and from instructions regarding criticality of instruments which are

administered in academic training. What are the results of these training procedures? Do pilots converge on an optimal scanning pattern for each maneuver over the course of training, or does each pilot adopt a unique cross-check? Conversations with pilots would certainly indicate the latter. Some pilots maintain that they are able to take in the entire instrument display with a single glance while others report the use of a variety of sequential scan patterns. It was the intent of this research to apply techniques used in visual scanning to the cross-check situation to objectively determine pilot strategy. Two experiments were conducted to this end.

In Experiment I, students and instructor pilots were shown slides of the T-37 flight instruments depicting straight and level unaccelerated flight. Some slides contained significant deviations from a predetermined course. Detection of errors and time to respond to errors were monitored in an attempt to measure cross-checking efficiency and strategy. In the present task errors were indicated on one of five flight instruments: AI, heading, vertical velocity indicator (VVI), airspeed, and altitude.

The use of reaction-time as a dependent variable stems from the subtraction method developed by Donders in the late 19th century (e.g., Woodworth and Schlosberg, 1954). Consider two tasks, the second of which requires all of the mental operations of the first plus one. According to the subtraction method, the duration of the nonoverlapping mental operation or stage can be determined by subtracting the reaction-time of task 1 from task 2. The subtraction method is based on two assumptions: first, that reaction-time is composed of a series of discrete stages, and second, that an additional stage, when added, doesn't interact with the other stages, i.e., the assumption of pure insertion. Methods for assessing the validity of these assumptions have been developed (e.g., Sternberg, 1969a, 1969b) and variants of the subtraction method have recently been applied to the study of visual scanning. These investigations have shown reaction time to be a reliable indicator of both serial (Sternberg, 1969a) and parallel (Neisser, 1967) processing. Thus, response latency has the flexibility to cover the range of scanning patterns which might be encountered in the current project.

How would the use of a cross-check procedure affect reaction-time? In the current project the display size (number of instruments) was held constant and the position of the error in the display was manipulated. If pilots use a fixed cross-check, the pattern of the cross-check

should be revealed by subtraction. If we assume that scanning proceeds in serial beginning with instrument 1 and ending with instrument N, which contains the target, then the response latency is the sum of the scan and decision times for the N-1 preceding instruments plus scan and decision time for the Nth instrument, plus a residual response time component. Thus reaction time will increase for targets presented on instruments scanned later in the scanning pattern. The pattern in which instruments are scanned can be determined by rank-ordering the response latencies according to which instrument contained the target.

Consider some possible differences in scanning between students and IPs which might be observed given hypothesized functions of training and experience in the aircraft. First, there is a possibility with a concentrated flight instrument system (such as exists in the T-37) that a pilot might learn to observe all the instruments in a single glance or he could process in parallel. If that were the case, there would be no differences in reaction time or accuracy as a function of the position of the error in the display. Second, assume that sequential patterns are employed and that there is an optimal scanning pattern for each maneuver. One function of training and experience might be a convergence toward that optimal pattern. Thus in terms of the present experiment, IPs as a group should show a more consistent trend toward a particular pattern than should students. Yet another possibility concerns the differential instructions given during training. Students are given examples of cross-checks for each maneuver, but they are also told to develop a cross-check which works best for them. If students followed the latter strategy, that would lead to highly consistent scanning patterns for individuals but not necessarily for groups. With their additional flight experience IPs should show the most consistent individual patterns, while students, still in the process of learning, would show variance resulting from trying out different patterns.

II. METHOD

Subjects

Two groups of twelve subjects were employed. One group consisted of students in the undergraduate pilot training (UPT) program at Williams Air Force Base, Arizona. All had previously seen the T-37 instrument panel and one had completed the UPT program. The second group consisted of experienced instructor pilots (IPs) from the UPT programs.

Apparatus and Stimuli

The experiment was performed in a 6 by 8 foot sound insulated room within the Flying Training Division, Air Force Human Resources Laboratory (AFHRL/FT) audio-visual laboratory at Williams Air Force Base. The apparatus consisted of a tachistoscope, rear projection screen, response levers, and an electronic interval timer (Hunter model 220c). The subjects were seated, individually, facing the rear projection stimulus presentation screen, which was rigidly mounted atop and perpendicular to the work surface of a small utility table. Located on the table and immediately in the front of the viewing screen were two firmly mounted response levers; the leftmost lever signalling a correct response and the rightmost lever an incorrect response. The experimenter sat adjacent to and facing the subject where the "incorrect/correct" response was monitored as well as the elapsed time from onset of stimulus to subject's response. A Kodak Carousel model 800 H, modified to tachistoscopic configuration via a Lafayette Electronics timed shutter, presented the visual stimulus materials for 2 seconds with 5 second interslide intervals.

Construction of Stimuli

A total of 920 stimulus slides were used. Half of these depicted a straight and level course with a heading of 270°, an airspeed of 160 kt, an altitude of 15,000 ft., and a vertical velocity of zero. The other half depicted a straight and level course with heading of 305°, airspeed of 164 kt, an altitude of 15,250 ft., and a vertical velocity of zero. Certain of the slides (targets) contained an error on one of the instruments, which was to be detected by the subject. A target could contain an error on any of the five instruments: heading, airspeed, altitude, VVI or attitude indicator. But on any given target slide only one instrument was shown in an error state.

The following procedure was used to define the errors: A computer printout of the performance of an advanced student flying straight and level in the T-4G simulator was obtained from AFHRL. This printout contained 60 samples obtained at a 1/sec. rate. The variance of each of 6 flight parameters (pitch, bank, heading, airspeed, altitude and vertical velocity) was calculated. Research pilots at AFHRL were interviewed to determine the minimum deviation in altitude that would be called an error. This turned out to be 100 feet. The Z score calculated for 100 feet was 7.3σ. Error magnitudes for the remaining 5 parameters were set at 7.3 times the standard deviation of that parameter to generate a "small error" condition. A "large error"

condition was obtained by multiplying the standard deviation of each parameter by 14.6.

Thus, considering the normal variance of an instrument in straight and level flight to be noise, in which the error target is to be detected, we can say that errors on all instruments were equated for detectability. The error magnitudes employed were: Pitch - 5^{o}; Bank - 10^{o}; Heading - 5^{o}; Airspeed - 11 kt; Altitude - 90 ft; VVI - 1000 ft/min; magnitudes of large errors were twice those of the small errors (see Figures 1 and 2). The same error magnitudes were employed for both courses.

Four trays of 115 slides were made for each course. Four error densities were used for each course: 60 errors in 115 slides, 50/115, 30/115, and 20/115.

Procedure

The subjects were seated and the physical layout of the experiment was briefly described, i.e., tachistoscope, screen, response levers, and interval timer. Additionally, in an attempt to put the subject at ease, a brief explanation of our interest in studying instrument cross-check techniques was briefly conveyed. Specific written instructions regarding the types of slides to be presented (coupled with an actual demonstration), what the instrument settings were designed to represent, the presentation and interslide durations, what constituted a correct and incorrect response, the fact that a given slide contained at most only one error, and instructions to be as accurate as possible and to respond upon detection of an error, were read to the subjects. The slide carousels were presented in random order and alternated between course conditions. Prior to the onset of the first slide from a given carousel, the nominal course settings as well as the course tolerances were orally presented to the subject. Values and limits were repeated to the subject as many times as requested. The task of the subject, once given course and course limits for a given carousel, was to view each stimulus presentation and to respond "correct" if the slide contained no errors and to respond "incorrect" if erroneous settings or indications outside of the allowable tolerances were detected. Two dependent measures were taken, target detection and response latency.

III. RESULTS

First it is important to determine if detection accuracy and response latency are sensitive to flight experience, thus providing some measure of the validity of our procedures.

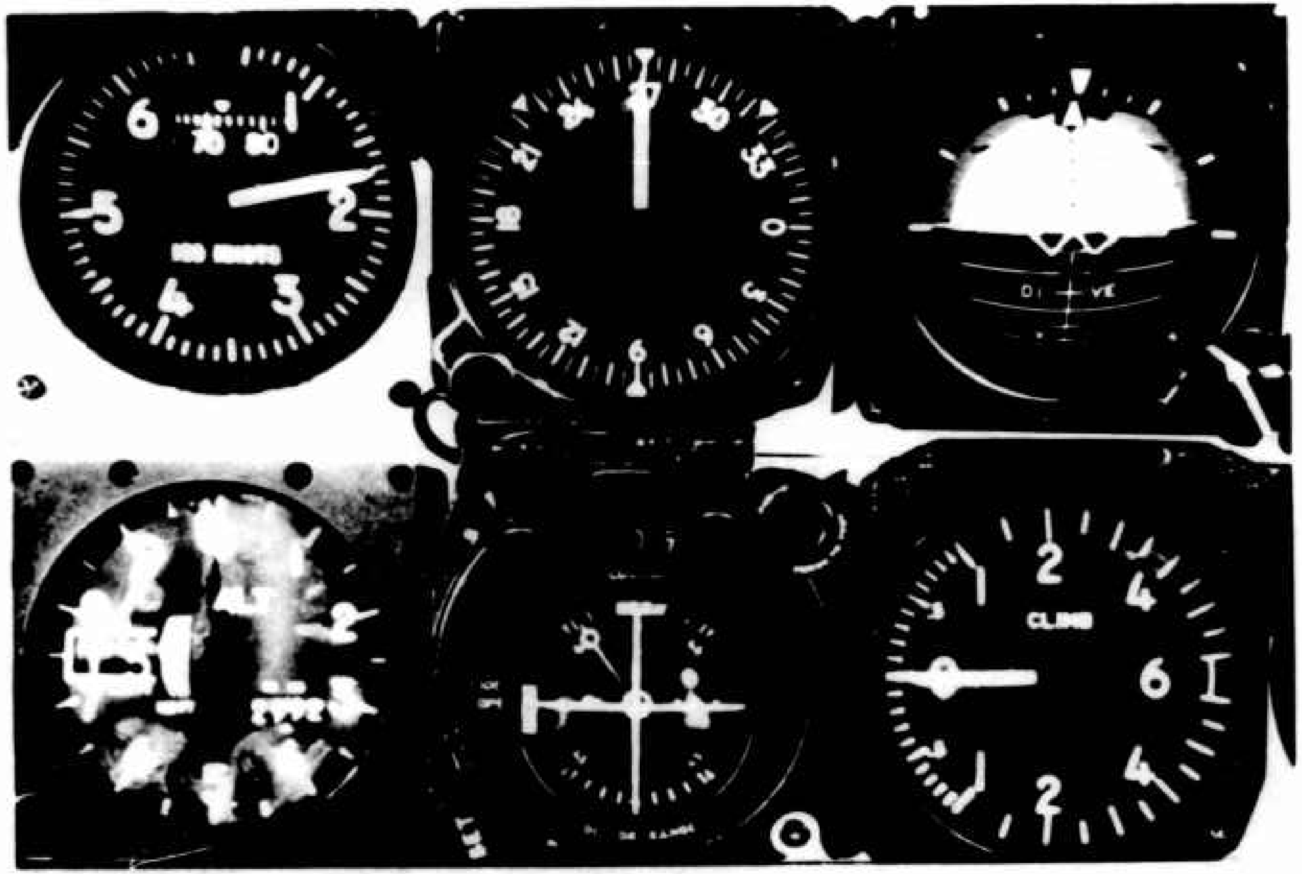

Figure 1: Stimulus display for Experiment I showing small airspeed error.

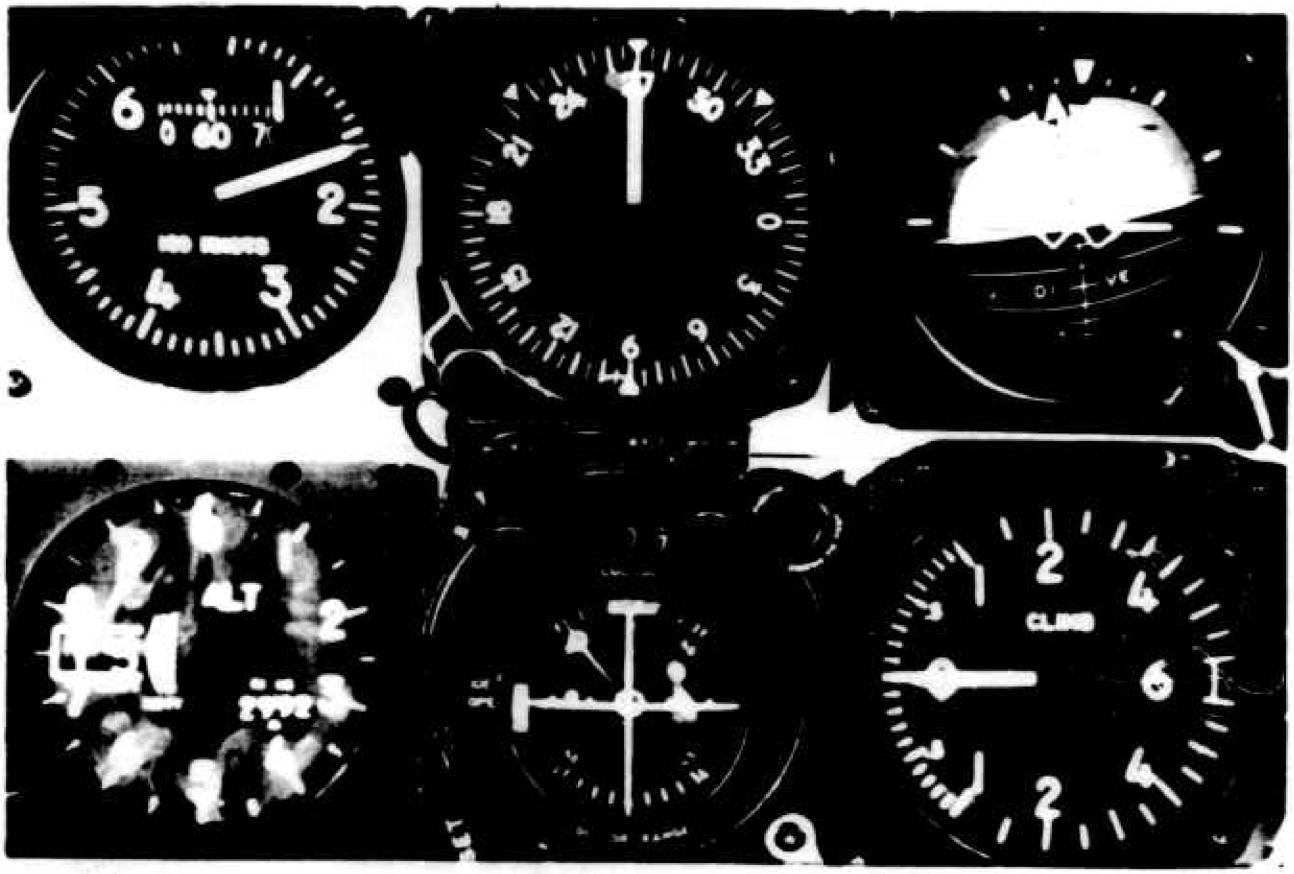

Figure 2: Stimulus display for Experiment I showing small bank error.

If accuracy of cross-checking is assumed to be directly related to the amount of flying experience, then it would be expected that IPs would detect targets with greater accuracy than would students. Using the theory of signal detectability methodology and converting subject's raw scores to appropriate hit (detection of target when target presented) and false alarm (detection of target when no target presented) rates, a value for d' was determined for each subject collapsed across all conditions. These data, presented in Table 1, and subjected to Mann-Whitney U tests, indicated as expected that IPs detected error targets with greater accuracy than did student pilots, $U(12,12) = 18$, $p < .001$, irrespective of the course, the size of the error target, or which instrument contained the error.

Assuming that speed of detection in cross-checking increases with increasing flight experience, it was expected that IPs would detect errors faster than student pilots. Collapsing the IP and student pilot response latency data across instruments and error conditions, i.e., large and small, and analyzing via Mann-Whitney U tests, significant differences between student pilot and IP reaction times were obtained, as expected. That is, IPs reacted faster in error detection than did student pilots, $U(12,12) = 22$, $p < .01$ (see Table 2).

Having shown the dependent variable sensitive to flight experience, response latencies aimed at revealing cross-check patterns were analyzed. It was assumed that systematic scan or cross-check patterns would be reflected in the cross-correlations of response latency, as a function of specific instrument error.

If RT to a target on instrument N is assumed to be the sum of N scan and decision times, then varying error magnitude should affect RT by altering the Nth decision time. This effect should be in the same direction, though not necessarily of the same magnitude, for all instruments, and should occur on the instrument actually containing the error. Thus while varying error magnitude should alter the absolute RT value, rank-orderings arising from the use of systematic scanning patterns should be preserved. That is, using Spearman rank-order correlations between large and small magnitude error latencies, a high correlation would indicate a consistent cross-check pattern while low correlation would indicate an inconsistent cross-check pattern. Spearman rank-order correlation coefficients were computed for the latency data of the 12 student pilots and the 12 IPs. Because of the large number of correlations being computed, a very strict criterion was applied in testing the significane of the correlation. Using an alpha of .002

TABLE 1

VALUES FOR d', TARGET DETECTION*

STUDENT PILOTS	INSTRUCTOR PILOTS
2.17	2.78
2.39	2.81
2.48	2.85
2.55	2.89
2.64	2.89
2.71	2.96
2.74	3.05
2.76	3.26
2.79	3.29
2.80	3.42
3.06	3.44
3.29	3.65

*$U(12,12) = 18$, $p < .001$

TABLE 2

MEAN VALUES OF RT, COLLAPSED ACROSS

INSTRUMENTS AND ERROR CONDITIONS*

STUDENT PILOTS	INSTRUCTOR PILOTS
1.49	1.39
1.57	1.60
1.69	1.64
1.83	1.66
1.83	1.69
1.95	1.74
2.00	1.75
2.06	1.80
2.23	1.80
2.24	1.86
2.29	1.91
2.71	2.38

*$U(12,12) = 22$, $p < .01$

for each individual test, an overall alpha of .05 was obtained for the experiment. With this criterion, four students and no IPs evidenced a consistent scanning pattern (see Table 3). According to a χ^2 test for two independent samples this difference was reliable ($\chi^2(1) = 4.8$, $p < .05$). These data suggest that in fact students show a tendency to use a consistent scanning pattern while IPs do not, a finding which is in direct opposition to our expectations.

Briefly summarizing, the results indicate: a) IPs detect errors with greater accuracy than do student pilots, b) IPs are faster at detecting errors than students, and c) systematic cross-check patterns did not appear to be employed by IPs while student pilots appeared to utilize systematic patterns.

IV. DISCUSSION

The present study investigated three aspects of instrument scanning performance: (a) accuracy, (b) latency to detect target conditions, and (c) the use of systematic scanning patterns. With regard to two of these aspects, accuracy and latency, experienced pilots showed superior performance to students. This superior performance obtained despite the fact that the IPs did not use any detectable scanning pattern. It had been assumed that systematic scanning would lead to superior performance; however, student pilots, who showed a systematic scanning pattern, showed both poorer detection accuracy and longer response latency. One possible explanation for this finding is that instrument reading is such an overlearned behavior for experienced pilots that they can take in the entire display at a single glance. This seems unlikely, however, since the mean latencies for eleven of the twelve IPs were within the range obtained for the students. A more likely explanation is that experienced pilots are able to adapt their scanning strategy to fit the scanning task. That is, rather than using a rigid scanning pattern, they use a flexible scanning strategy which allows them to emphasize important or difficult aspects of the display.

The notion of a flexible scanning pattern should not be viewed as antithetical to that of a fixed cross-check, but rather flexible scanning should be viewed as a complementary process. In early training the student is instructed in the fixed cross-check procedure. He is also encouraged as he progresses through the program to adapt this procedure to meet his own needs. The present study has shown that in fact experienced pilots seem to rely more

TABLE 3

SPEARMAN RANK-ORDER CORRELATIONS BETWEEN SMALL AND LARGE MAGNITUDE ERROR LATENCIES. ASTERISKS INDICATE STATISTICALLY SIGNIFICANT CORRELATIONS ($p < .001$)

STUDENT PILOTS (R_T)	INSTRUCTOR PILOTS (R_T)
.94	.40
.90	.60
1.00 *	.49
1.00 *	.77
.94	.94
.54	.89
1.00 *	.94
1.00 *	.89
.94	.83
.83	.77
.71	.66
.60	.37

heavily on their own individual, adaptive scanning strategies than they do on the more rigid strategies taught them in early training.

Experiment II concentrated on the individual adaptive aspects of a pilot's scanning strategy. No attempt was made to define these aspects since they doubtlessly differ from pilot to pilot and task to task. Rather, Experiment II concentrated on the generalizability of such strategies between tasks and their effects on task performance.

V. Experiment II

In Experiment I, consistent scanning was defined to be using exclusively a single scanning pattern for the particular display scanned. Implicit in this definition is the notion that there exists a particular pattern, the use of which will result in optimal scanning performance. The display used in Experiment I was amenable to such a definition for several reasons. The first and most important reason is that there was little need to recheck an instrument once it had been scanned. This was so because the targets (errors) were highly detectable; d' in all cases was well over 2 and in many cases over 3, and because the display was static, that is, instrument readings did not change after the instruments had been checked. In this regard this task differed from the actual flight situation. In the flight situation errors may occur simultaneously on multiple instruments. Thus, the difficulty of the detection task is increased, and so the need to recheck an instrument is increased. Also in the flight situation, a reading can change immediately after the instrument is checked. Thus, in the experimental situation the subject could be more certain that his reading was correct without rechecking the instrument than could a pilot in an actual flight situation. If in fact experienced pilots rechecked the instruments before responding out of habits acquired in flight, they would fail to show a consistent scanning pattern while the student, who had not developed such habits and based his judgement on only one scan, would show the optimal scanning pattern.

The experimental task also differs from actual flight in that the probability and importance of errors on any instrument is constant over time. In the flight situation both of these factors vary between maneuvers as well as varying as the pilot progresses through a maneuver. This forces the pilot to vary the emphasis he places on specific instruments over time. If this practice were employed in the experimental task, the data when averaged would fail to demonstrate consistent scanning although on any given trial

the pilot might be scanning consistently.

The differences between the requirements of the task of Experiment I and the flight situation can be summed up as the difference between a static scanning situation and a dynamic scanning situation. A task which is to determine how pilots behave in the flight situation should be a dynamic task. That is, the nature of the targets in given areas of the display should all vary over time. In the dynamic type of scanning task the single, optimum scan definition of consistency used in Experiment I is inappropriate. Since the demands of the task vary over time, there is no single, optimum scanning pattern. Rather the task of the observer is to adapt his scanning strategy to meet the changing requirements of the task without allowing his attention to lapse. That is, in the dynamic task consistent scanning is best defined in terms of low distractability from the scanning task.

One technique which has been used to measure distractability is the Attention Diagnostic Method (ADM) (Block, 1975). This task has been employed in a variety of industrial settings to predict rates of accidents resulting from failures of attention. Basically this task consists of a 5 by 10 matrix containing the numbers 10 through 59 in random sequence. The rows of the matrix are presented in different colors (Block used the colors red, blue, yellow and green). Subjects are asked to scan the matrix finding each number in sequence and reporting its color. The latency of each response is measured. Distractability is indicated by the incidence of long latency responses. The present experiment employed the ADM to investigate the distractability of experienced pilots (IPs), student pilots from UPT, and Arizona State University (ASU) undergraduates. The IP group was expected to demonstrate less distractability in scanning--that is, fewer long latency responses than student pilots and college students.

VI. METHOD

Subjects. Three groups of subjects were employed. The instructor pilot (IP) group consisted of 9 T-37 instructors from Williams Air Force Base. The student pilot (SP) group consisted of 9 students from the Williams Air Force Base undergraduate pilot training (UPT) Program. Flight experience in this group ranged from no jet experience for 2 students to approximately 200 hours of jet time for 2 students who had completed the program but had not yet been reassigned. The third group (ASU) consisted of 9 male students from an introductory psychology class. None had

flight experience in any type of aircraft.

Apparatus. Stimuli were rear projected using a Kodak 800L carousel slide projector. Latency of subject's responses was measured with a Hunter model 220C counter to the nearest second. Subject's responses were made on two micro switches mounted on the table before the subject.

Stimuli consisted of four 10 by 10 matrices containing the numbers 10 through 59 (Figure 3). The numbers were 5/8 inch in height. The matrix was divided into fifty one inch rectangles, and each rectangle contained one number. Each matrix contained ten rows of five numbers. The numbers in half of the rows in each matrix were red and in the other half the numbers were white with colors alternating between rows. The numbers were assigned positions in the matrix as follows: First a prototype matrix was constructed. Two such prototypes were used. In one, the top row was red, and in the other, the top row was white. Numbers were then assigned randomly to positions in the matrix beginning in the upper left corner and working left to right and down. Each of the numbers 10-59 appeared once in the matrix. Subsequent matrices were generated by randomly moving half of the numbers in the prototype to a new position in a row having the same color as the original row. Thus two sets of matrices were generated. Within each set the colors of the numbers were constant, but the positions of the numbers varied.

Procedure. The subject was seated in a dimly lit room approximately 2 feet from the screen. He was shown a matrix and told that it was typical of the matrices he would see in the task. He was told that as soon as the next slide appeared, he was to begin searching immediately for the number 10. As soon as he found it, he was to press the left button if it were red and the right if it were white. He was to follow the same procedure for the number 11 and so on until he had completed all fifty numbers. It was explained that the slide would remain on until the entire task had been completed. He was told to work as rapidly as possible since the cumulative latency of each response from the onset of the slide was being measured. However, accuracy was stressed and he was told not to guess although if he were certain that he remembered the color of a number from having seen it during an earlier search, he need not find the number a second time.

The subject was run through four matrices without feedback regarding the latency of his responses. He was given a break of at least 30 minutes after which he would

16	45	33	54	10	Red
34	26	28	36	14	White
44	37	41	56	40	Red
15	38	50	52	57	White
20	47	31	22	32	Red
24	17	12	35	49	White
43	39	42	59	23	Red
19	58	46	29	11	White
30	48	13	51	25	Red
55	27	53	21	18	White

FIGURE 3

Visual Display, Attention Diagnostic Method, Experiment 2

repeat the same four matrices. Two matrices were based on the white prototype and two were based on the red prototype. Questioning after the end of the second session revealed that subjects were not aware of the similarity of slides based on the same prototype, nor were they aware that the same four slides were used in both sessions.

Results and Discussion. It was expected that the IP group would show superior initial performance since they would have a lesser tendency to be distracted from their scanning pattern. Figure 4 shows the mean latencies over trials for each of the three groups to complete scanning for all fifty numbers. The prediction of initial superiority was not supported. In fact, on trial 1 the mean latency of the IP group was greater than that of either the SP or ASU groups although none of the differences reached significance.

If pilots are not superior at novel scanning tasks, the next question is this: Do they show greater facility at learning new scanning tasks? When an observer is introduced to a novel scanning task, his initial scanning strategy may not be the most efficient. The pattern of his eye movements may not maximize his probability of spotting the target and his cognitive processing strategy may lead him often to attend to irrelevant aspects of the display at the expense of more relevant aspects. This is to say he will fail to attend to relevant aspects of the display and will frequently "misdirect" his attention. The result of this misdirection of attention in the ADM task will be a high incidence of long latency responses. As scanning becomes more efficient with practice, two changes can occur which would decrease response latency. First is that the observer can modify his scanning procedure to avoid failures to fixate on some parts of the display. For instance, he might pick certain reference points in the display to insure that his eyes pass over all parts of the display, or he might adjust his rate of scanning to minimize failures of attention while at the same time maintaining a reasonable modal scan-time. At a cognitive level, certain features of the display might be given higher priority. For example, in the ADM task the second digit of each pair (e.g., the 1 in 51) carries more information than does the first. Attending to the second digit of a pair first runs counter to our usual reading practices; however, with practice an observer might adopt such a cognitive strategy to reduce his scanning time. The adoption of improved scanning and cognitive strategies should have two effects on task performance. It should decrease the latency for each response and so the latency to scan the entire matrix because of the increase in scanning efficiency. Second, the incidence of

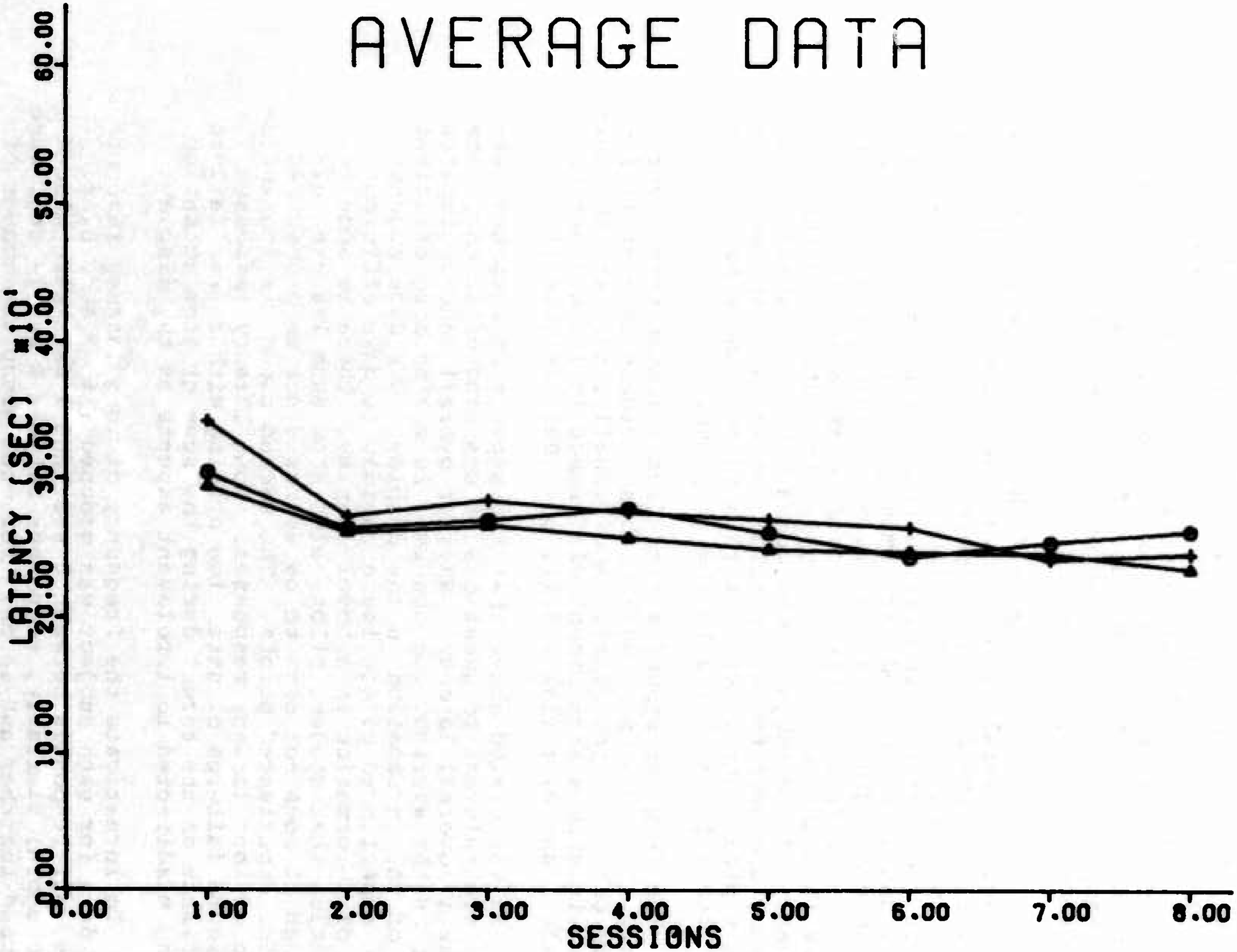

Figure 4: Mean latency to complete ADM task over trials for student and instructor pilots and ASU students.

long latency responses should be decreased through the reduction in the number of failures or misdirections of attention. It should be pointed out that while these two are similar, they are not identical. A decrease in the number of long latency responses will result in a decreased overall latency. However, if the observer were to reduce the frequency of long latency responses through slower, more meticulous scanning rather than more efficient scanning, the frequency of long latency responses might decrease while the overall latency increases.

It can be seen in Figure 4 that all three groups show decreases in overall latency over the course of eight trials. However, for the ASU group (see Figure 5), this difference fails to reach significance ($r = -.21$, $p > .05$). A comparison of the overall latencies on trials 1 and 8 also reveals no evidence of improved performance, $t(16) = 1.63$, $p > .1$. The SP group did show a significant linear trend over trials ($r = -.27$, $p > .05$). However, examination of Figure 6 shows that this effect is largely attributable to the performance of one subject, who began with an unusually long latency and over the eight trials worked his way down to a more typical score. Without this subject there is no evidence for a linear trend in the SP group.

The IP group showed a significant linear trend over trials ($r = .375$, $p < .001$) with a decrease in overall latency of 9 sec. per trial. Additionally, this was the only group to show a significant difference in latency between trials 1 and 8, $t(16) = 3.11$, $p < .005$ (Figure 7).

It was stated above that two aspects of response latency are relevant to questions of scanning efficiency. The first is overall latency. Shorter overall latency implies that on the average, the observer is making more efficient use of the information in the display. Our data suggest that experienced pilots learn rapidly to make efficient use of information in a novel display. There is some indication that student pilots also show such learning although it does not seem to be as rapid nor as pronounced as for experienced pilots. The second aspect is the incidence of long latency responses. Long latency responses indicate failures of attention arising either from failure to fixate on the target during the scan or from attention being misdirected to irrelevant aspects of the display.

To investigate the frequency of long latency responses, the data for each subject was grouped into 7 sec. bins. Thus, all responses between 0 sec. and 7 sec. were placed in one bin; similarly, responses between 8 and 14 sec. were grouped together and so on. For each group the number of

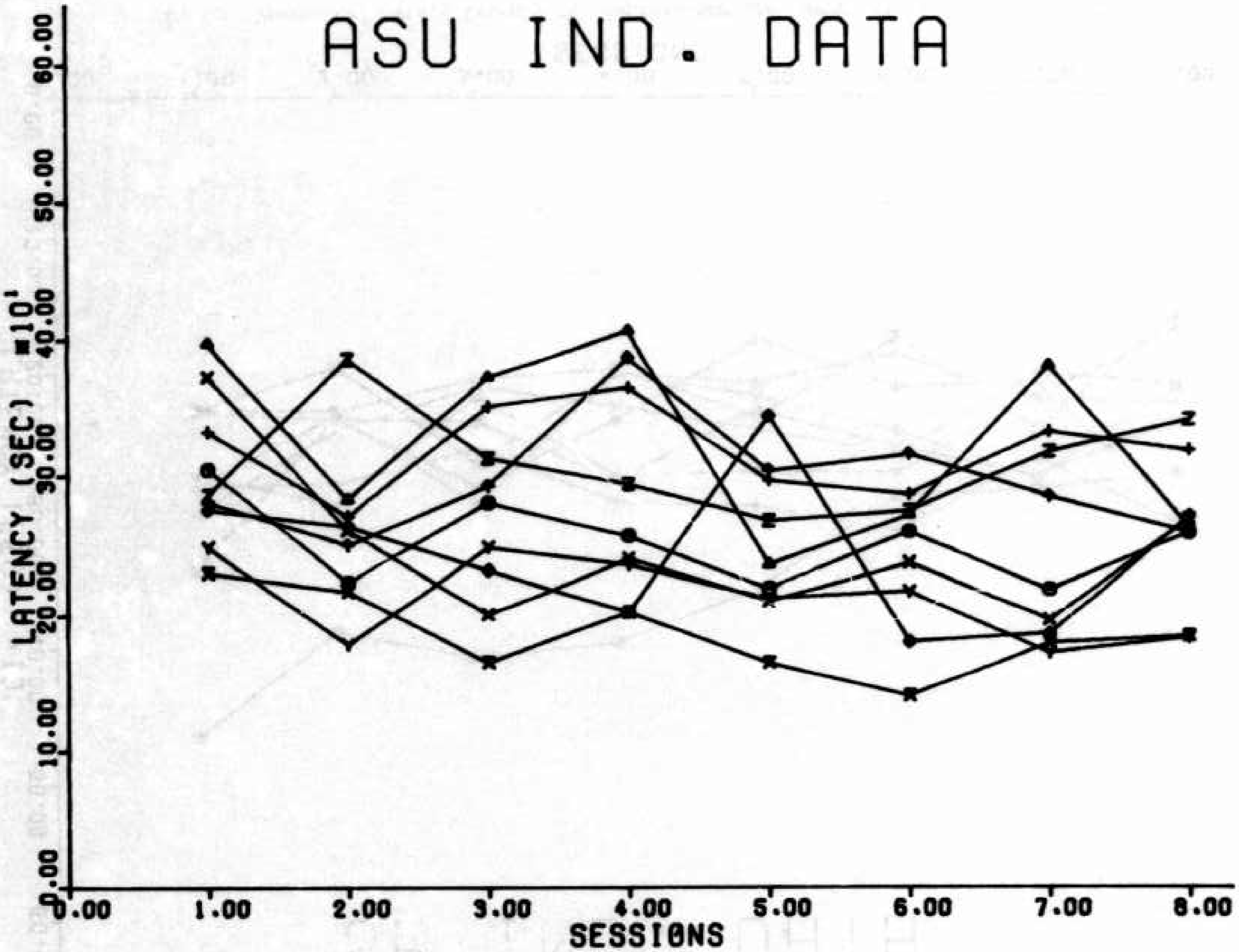

Figure 5: Individual data. Latency to complete ADM task over trials for ASU students.

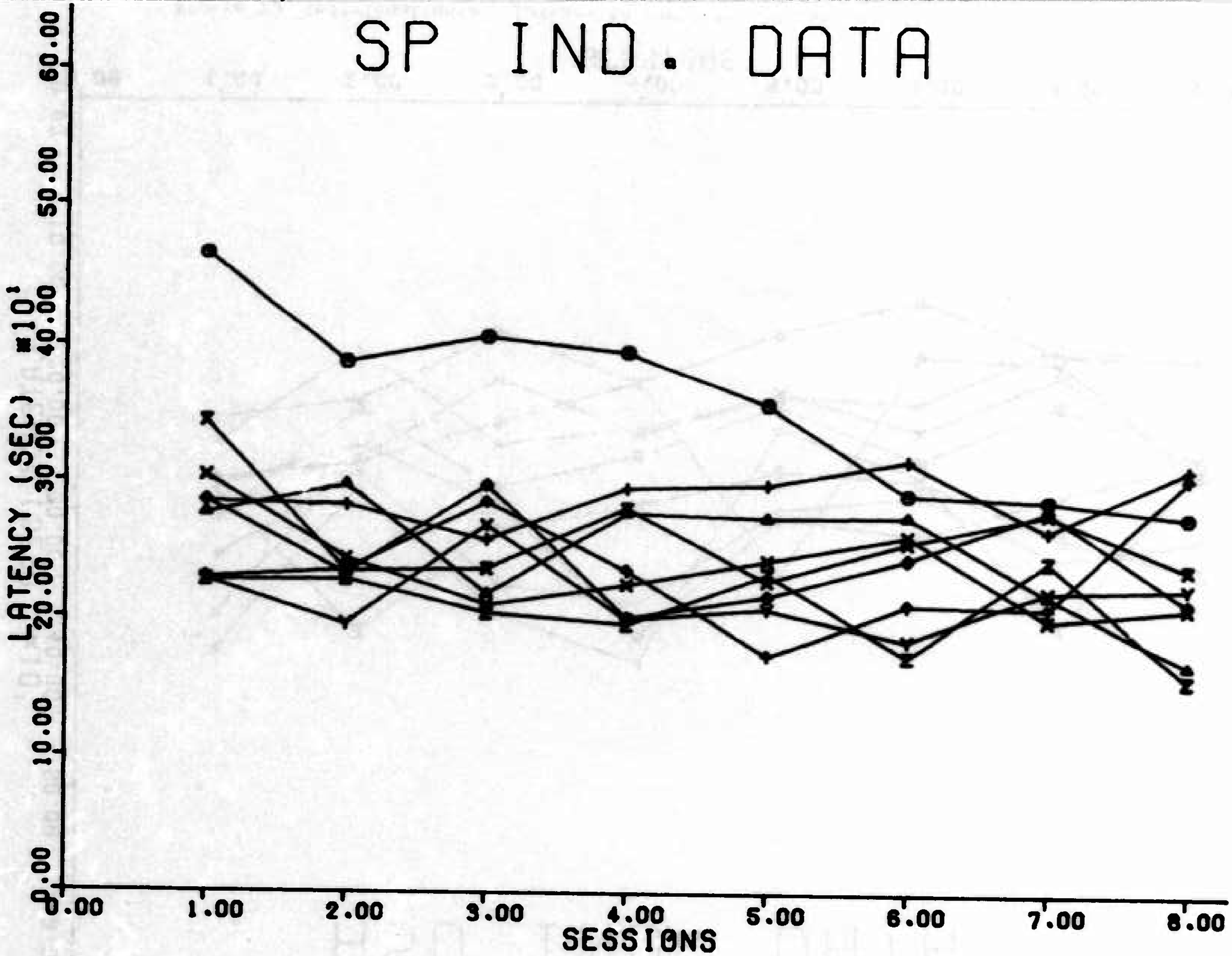

Figure 6: Individual data. Latency to complete ADM task over trials for student pilots.

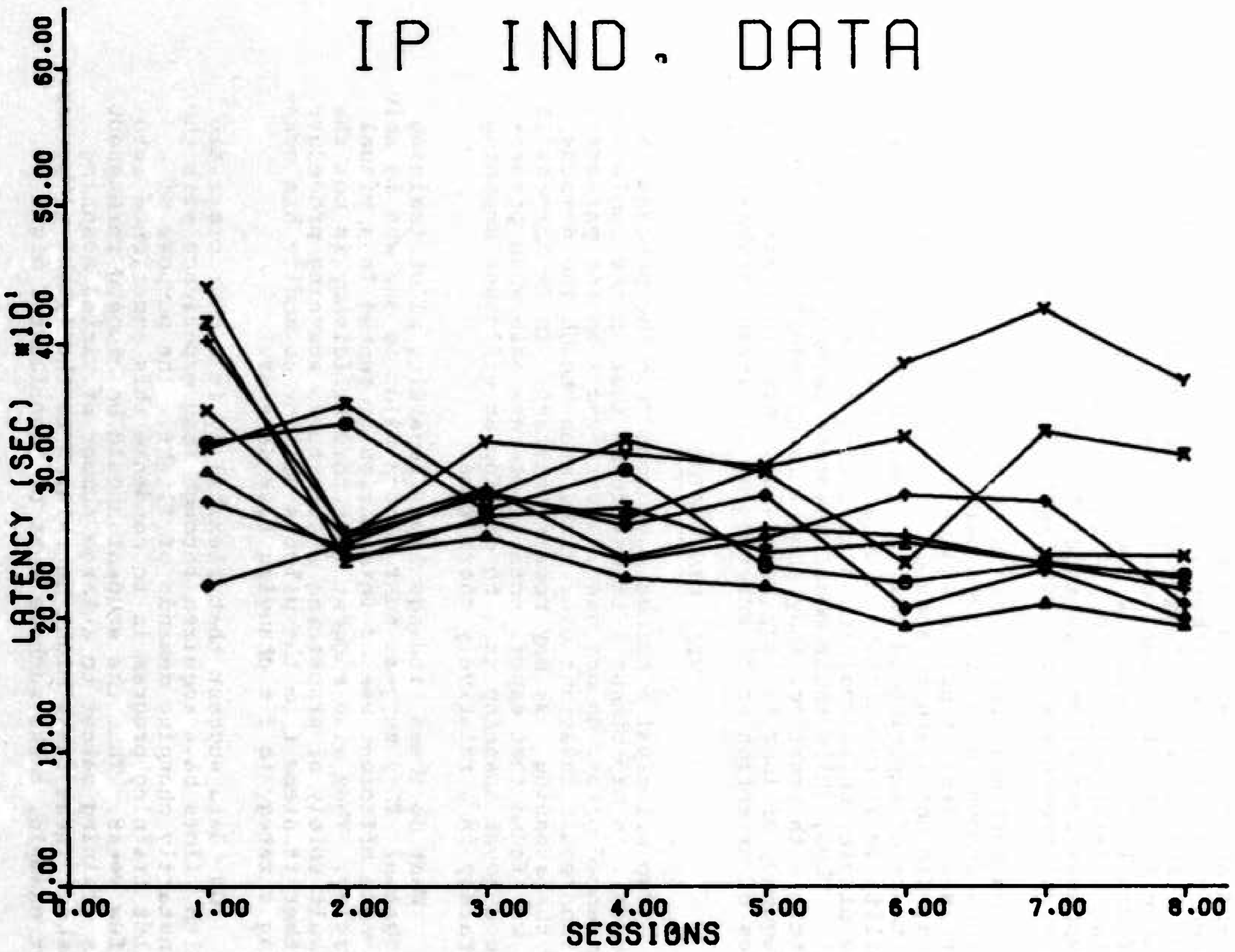

Figure 7: Individual data. Latency to complete ADM task over trials for instructor pilots.

responses over 7 sec. and over 14 sec. was determined for trials 1 and 8 by adding the number of responses in the bins greater than 7 sec. and 14 sec., respectively. The SP and ASU groups showed no significant differences between trials 1 and 8 either in the frequency of responses over 7 sec. or in the frequency of responses over 14 sec. (Table 4). The IP group, on the other hand, showed significantly fewer responses over 7 sec. on trial 8 than on trial 1 ($t(5) = 4$, $p < .01$). This group also showed fewer responses over 14 sec. on trial 8 than on trial 1 although this difference was not significant ($t(5) = 1.87$, $p < .1$).

On both measures of scanning efficiency experienced pilots demonstrated improved performance over the course of 8 trials. Student pilots and naive observers, on the other hand, did not learn to scan efficiently. The improved performance of experienced pilots seems to result from their ability to learn quickly to attend to relevant aspects of the display while avoiding distraction by irrelevant stimuli. Thus, while their modal response latency changes little with practice, they show a large reduction in the frequency of long latency responses, which indicate failures of attention or distraction by irrelevant stimuli.

VII. DISCUSSION

Two principal findings have come from the present research. In Experiment I it was found that on the whole experienced pilots do not use standardized scanning patterns, which remain invariant over time, even though the demands of the scanning task may remain constant. In Experiment II it was found that experienced pilots are able when presented a novel scanning task to develop an efficient scanning strategy in a relatively short time.

What do these findings imply regarding pilot training programs? They suggest that a good pilot is one who is able to make efficient use of information presented in a visual display. They also suggest that this efficiency is not the result solely of practicing a structured scanning procedure. Rather it stems from the pilot's ability to modify his scanning strategy to fit changing task demands.

Our data suggest that flexibility is a characteristic which pilots have acquired through long experience with the constantly changing demands of flight. The purpose of a pilot training program is to condense this experience into a few weeks. Thus the student should be exposed throughout his training career to a large number of varied scanning tasks. These tasks might be related to the flight situation; for example, tasks such as that of Experiment I might be

TABLE 4

	TRIAL					
		1			8	
Latency (sec) Subject	0-7	8-14	15+	0-7	8-14	15+
P1	24	13	12	31	15	3
P2	28	18	3	41	6	2
P3	22	21	6	30	18	1
P4	35	23	4	33	15	1
P5	33	11	3	34	12	3
P6	27	21	1	37	10	2
$\bar{X}$P	28	17	5	34	13	2
S1	32	13	4	37	7	5
S2	33	14	2	26	18	5
S3	20	20	9	23	26	0
S4	26	20	3	40	8	0
S5	27	16	6	34	14	1
S6	29	17	3	26	18	5
$\bar{X}$S	28	17	5	31	15	3
A1	28	15	6	30	17	2
A2	23	18	8	32	12	5
A3	29	12	8	27	18	4
A4	26	11	6	30	15	4
A5	32	15	2	37	10	2
A6	31	15	3	33	8	7
A7	32	14	3	26	17	6
A8	25	14	10	34	10	5
A9	36	10	3	40	7	2
$\bar{X}$A	29	14	5	32	13	4

employed using a variety of maneuvers and different instrument displays or they might be unrelated, like the ADM task. By increasing the student's exposure to novel scanning situations, such a program will teach the student flexibility in his approach to the scanning task and prepare him for the constantly changing demands of the actual flight situation.

REFERENCES

Block, J.R. Some preliminary results in predicting accidents with the attention diagnostic method. Progress report submitted to the Insurance Company of North America: MEND Institute, 1975.

Neisser, U. Cognitive Psychology, Appleton-Century-Crofts, New York: 1967.

Sternberg, S. Memory Scanning: Mental Process Revealed by Reaction-time Experiments: American Scientist, 57 (4), 1969a, 421-457.

Sternberg, S. The discovery of processing stages: Extensions of Donders' method. Acta Psychologica 30 Attention and Human Performance II, W.G. Koster (Ed.), 1969b, 276-315.

U.S. Air Force Flying Training Manual. Instrument Flying, AF Manual 51-37 (1971)

U.S. Air Force Flying Training Manual. Primary Flying, Jet, ATC Manual 51-4 (1974)

Woodworth, R.S. and Schlosberg, H. Experimental Psychology (Revised Edition), Henry Holt & Company, New York: 1954.

★U.S. GOVERNMENT PRINTING OFFICE: 1976- 671-613/16